

# A GAUSSIAN EXPONENTIAL APPROXIMATION TO SOME COMPOUND POISSON DISTRIBUTIONS

BY

WERNER HÜRLIMANN

## ABSTRACT

A three parameter Gaussian exponential approximation to some compound Poisson distributions is considered. It is constructed by specifying the reciprocal of the mean excess function as a linear affine function below some threshold and a positive constant above this threshold. As an analytical approximation to compound Poisson distributions, it is only feasible either for a limited range of the Poisson parameter or for higher coefficients of variation. A semiparametric determination of the unknown threshold parameter is proposed. The analysis of a real-life example from pension fund mathematics displays an improved quality of fit of the new model when compared with other simple good alternative approximations based on the zero gamma, translated gamma and zero translated gamma.

## KEY WORDS

Compound Poisson distribution, Gauss density, exponential density, mean excess function, stop-loss transform, analytical approximation, threshold parameter.

## 1. INTRODUCTION

Compound Poisson distributions define one of the most important class of distributions in applied probability and operations research. In practical work they are used in numerous fields such as insurance, finance, reliability and queueing theory. They are identified as the distributions of random variables of the type  $X = Y_1 + \dots + Y_N$  where  $N$  is  $\text{Poisson}(\lambda)$  and the severities  $Y_k$  are independent and identically distributed non-negative random variables, which have the same distribution as the random variable  $Y$  and are independent from  $N$ .

Roughly speaking there are three categories of methods for the numerical and statistical evaluation of these distributions. After a discretization of  $Y$ , it is possible to compute quickly close *numerical approximations* using the recursive method of Panjer (1981). Simple analytical approximations of the tails of

the distribution are obtained with the *asymptotic normal-power approximation* (e.g. Beard et al. (1984), Section 3.11), provided the skewness parameter is sufficiently small. Among the simplest *analytical approximations*, one finds often quite good approximations with gamma type distributions including the translated gamma, zero gamma and zero translated gamma distributions (e.g. Dufresne and Niederhauser(1997)).

Belonging to the latter category, a general reliability constructive approach proposes to specify first a higher degree stop-loss rate function (hazard rate, reciprocal of mean residual life and higher order generalization), then derive a higher degree stop-loss transform model, and through successive differentiation identify the desired distribution, as explained and illustrated in Hürliemann(2000b). Choosing a *linear affine function*  $\alpha + \beta \cdot x$  for the first three higher degree stop-loss rates generates three different Gauss type families of distributions, the first one being the Gaussian survival model by Kodlin(1967). The second one, which we consider in the present paper, defines the so-called *Gaussian stop-loss model*, whose stop-loss transform is of the type  $\pi(d) = E[(X - d)_+] = \mu \cdot \exp(-\alpha \cdot x - \frac{1}{2} \beta \cdot x^2)$ , with  $\mu = E[X]$  the mean.

Used as an analytical approximation to compound Poisson distributions, the Gaussian stop-loss model is only feasible either for limited „small” values of the Poisson parameter or for „high” values of the coefficient of variation. Albeit this unpleasant restriction, the present model might be useful in some situations involving either „rare events” or „high volatility”. Empirical comparisons show that the Gaussian stop-loss values  $\pi(d)$  approximate well for „medium” values of  $d$  around the mean. To improve the empirical fit for an extended range of higher values of  $d$ , especially in the right tail of the distribution, we assume that the reciprocal of the mean residual life function is a positive constant (in contrast to the Gaussian case for which this vanishes) above some „threshold” to be determined. This simplified assumption implies in particular an exponential tail distribution, which for some compound Poisson distributions is justified under a regularity condition (see Panjer and Willmot (1992), Section 10.2). A more detailed outline of the paper follows.

Some required prerequisites on reliability theory are summarized in Section 2. The details of the constructed Gaussian exponential stop-loss model are presented in Section 3. The conditions under which this model defines a feasible compound Poisson approximation are derived in Section 4. A semiparametric determination of the unknown threshold parameter of the model is proposed in Section 5. Finally, in Section 6, a real-life example by Held(1982) from the field of pension fund mathematics illustrates and justifies the use of the present model. Based on several significant goodness-of-fit measures, an improved quality of fit of the Gaussian exponential stop-loss model is found when compared with the simple alternative approximations based on the zero gamma, translated gamma and zero translated gamma distributions considered in Dufresne and Niederhauser(1997). We also analyse the performance of the Gaussian exponential approximation by varying the individual claim size distribution. The Gaussian exponential approximation fits better for a light tailed than a heavy tailed claim size distribution. As a remedy, we expect to see in future work a similar construction of the Gaussian Pareto model.

## 2. ON THE MONOTONE RELIABILITY CLASSES IFR AND DMRL

We recall the definitions of the monotone reliability classes IFR and DMRL. For any random variables  $X$  taking values in  $[0, \infty)$  with absolutely continuous distribution  $F(x) = \int_0^x f(t)dt$ ,  $f(x)$ , the density,  $\bar{F}(x) = 1 - F(x)$  the survival function, let

$$(2.1) \quad h(x) = -\frac{d}{dx} \ln \{\bar{F}(x)\} = \frac{f(x)}{\bar{F}(x)}, \quad x > 0, \text{ the failure rate,}$$

$$(2.2) \quad m(x) = E[X - x | X > x] = \frac{\pi(x)}{\bar{F}(x)}, \quad x > 0, \text{ the (conditional) mean remaining life function,}$$

where  $\pi(x) = \int_x^\infty \bar{F}(t)dt$  is called the *stop-loss transform*.

A given function  $m(x) : (0, \infty) \rightarrow (0, \infty)$  is the mean residual life function of some  $X$  if, and only if, it is right-continuous,  $0 < m(0^+) < \infty$ ,  $\int_0^\infty \frac{dx}{m(x)} = \infty$ , and  $H(x) = \int_0^x \frac{dx}{m(t)} + \ln \{m(x)\}$  is non-decreasing on  $(0, \infty)$  (see Bhattacharjee(1982)). The corresponding survival function is uniquely determined by the inversion formula

$$(2.3) \quad \bar{F}(x) = \bar{F}(0^+) \cdot \frac{m(0^+)}{m(x)} \cdot \exp \left\{ -\int_0^x \frac{dt}{m(t)} \right\}, \quad m(0^+) = \frac{E[X]}{\bar{F}(0^+)}.$$

A random variable  $X$  is of *class IFR (DFR)* if  $h(x)$  is increasing (decreasing) in  $x > 0$ , and it is of *class DMRL (IMRL)* if  $m(x)$  is decreasing (increasing) in  $x > 0$ . One knows that  $\text{IFR} \Rightarrow \text{DMRL}$ , but the converse does not hold (Bryson and Siddiqui(1969), an explicit counterexample is found in Muth(1977)). However, the converse  $\text{DMRL} \Rightarrow \text{IFR}$  is fulfilled provided the following inequality holds (Hürlimann(2000a), special case of Theorem 3.2):

$$(2.4) \quad m'(x) \leq m(x)^2 \cdot \frac{d^2}{dx^2} \ln \{m(x)\}, \quad \text{for all } x > 0.$$

This condition is implicit in Muth(1977), and it follows by taking the derivative of the well-known identity  $h(x) = \frac{1 + m'(x)}{m(x)}$  and making some rearrangements.

In particular, if  $m(x)$  is convex decreasing (Muth(1977)), or log-convex decreasing (Hürlimann(2000a)), then  $X$  is of class IFR. Common distributions, which are both DMRL (IMRL) and IFR (DFR), include the truncated normal, the Gamma and the Weibull (Barlow and Proschan(1975)), the Pareto of types II and III (see Davis and Feldstein(1979)), and the Benktander of types I and II (see Benktander(1970), Beard et al.(1984), Embrechts et al.(1997), Hürlimann(2001)). A computational criterion to test whenever a discrete arithmetic distribution is of class IFR or/and DMRL is formulated in Theorem 2.1 below.

Let  $X$  be a discrete arithmetic random variable with probabilities  $f_n$ ,  $n = 0, 1, 2, \dots$  such that  $f_0 > 0$ . Appropriate definitions for the discrete case are obtained from the continuous one by replacing functions  $f(x)$ ,  $x \geq 0$ , by sequences  $f_n$ ,  $n = 0, 1, 2, \dots$ . Given the probabilities, the survival sequence is defined by the recursion

$$(2.5) \quad \bar{F}_{n+1} = \bar{F}_n - f_{n+1}, \quad \bar{F}_0 = 1 - f_0, \quad n = 0, 1, 2, \dots$$

Then the stop-loss transform satisfies the recursion

$$(2.6) \quad \pi_{n+1} = \pi_n - \bar{F}_n, \quad \pi_0 = \mu, \quad n = 0, 1, 2, \dots$$

The failure rate and the mean residual life sequences are given by

$$(2.7) \quad h_n = \frac{f_n}{\bar{F}_n}, \quad m_n = \frac{\pi_n}{\bar{F}_n}, \quad n = 0, 1, 2, \dots$$

We derive a simple criterion for  $X$  to be of class IFR or/and DMRL.

**Theorem 2.1.** A discrete arithmetic random variable  $X$  with finite mean  $\mu > 0$  and probabilities  $\{f_n\}$ ,  $f_0 > 0$  is of class IFR if, and only if, the survival sequence  $\{\bar{F}_n\}$  is logarithmically concave, that is  $\nabla^2 \ln\{\bar{F}_n\} = \ln\{\bar{F}_{n+1}\} - 2 \cdot \ln\{\bar{F}_n\} + \ln\{\bar{F}_{n-1}\} \leq 0$ , or equivalently

$$(2.8) \quad (\bar{F}_n)^2 \geq \bar{F}_{n-1} \cdot \bar{F}_{n+1}, \quad n = 1, 2, \dots$$

Similarly,  $X$  is of class DMRL if, and only if, the stop-loss transform sequence  $\{\pi_n\}$  is logarithmically concave, that is

$$(2.9) \quad (\pi_n)^2 \geq \pi_{n-1} \cdot \pi_{n+1}, \quad n = 1, 2, \dots$$

**Proof.** The condition (2.8) is seen equivalent to  $h_{n-1} \leq h_n$  by making use of the recursion (2.5). Taking logarithms, this is equivalent to  $\{\bar{F}_n\}$  being logarithmically concave. The statement (2.9) is similarly shown equivalent to  $m_n \leq m_{n-1}$  by using the recursion (2.6).  $\diamond$

### 3. CONSTRUCTION OF THE GAUSSIAN EXPONENTIAL STOP-LOSS MODEL

Our choice of the mean residual life function is based on the following observations. For “medium” values of  $x$  around the mean, the reciprocal of  $m(x)$  for a discrete arithmetic compound Poisson distribution, which is the prototype of the aggregate claims model in classical risk theory, is approximately linear provided the Poisson parameter is sufficiently small. For “high” values of  $x$ , one can assume that  $m(x)$  is approximately constant under a regularity condition

(see Panjer and Willmot(1992), Section 10.2). Applying the described criterion, one obtains the feasible model

$$(3.1) \quad m(x) = \begin{cases} \frac{1}{\alpha + \beta x}, & 0 \leq x \leq x_0 \\ \frac{1}{\alpha + \beta x_0}, & x \geq x_0, \end{cases}$$

$x_0$  the threshold, with the parameter restriction

$$(3.2) \quad 0 < \beta \leq \alpha^2,$$

which follows from the condition that the corresponding density function must be non-negative. Since  $m(x)$  is convex decreasing, the corresponding random variable belongs to both classes DMRL and IFR. From (2.3) one obtains through application of a mean scaled reparametrization, after some straightforward calculations, the survival function

$$(3.3) \quad \bar{F}(x) = \begin{cases} \left\{ \alpha + \left( \frac{\alpha}{\gamma} \right)^2 \cdot \frac{x}{\mu} \right\} \cdot \exp \left\{ -\alpha \cdot \frac{x}{\mu} - \frac{1}{2} \left( \frac{\alpha}{\gamma} \cdot \frac{x}{\mu} \right)^2 \right\}, & 0 \leq x \leq x_0, \\ \bar{F}(x_0) \cdot \exp \left\{ -\left( \frac{x - x_0}{m(x_0)} \right) \right\}, & x \geq x_0, \end{cases}$$

and the stop-loss transform

$$(3.4) \quad \pi(x) = \begin{cases} \mu \cdot \exp \left\{ -\alpha \cdot \frac{x}{\mu} - \frac{1}{2} \left( \frac{\alpha}{\gamma} \cdot \frac{x}{\mu} \right)^2 \right\}, & 0 \leq x \leq x_0, \\ \pi(x_0) \cdot \exp \left\{ -\left( \frac{x - x_0}{m(x_0)} \right) \right\}, & x \geq x_0, \end{cases}$$

where  $\mu = E[X] > 0$  is the mean,  $\alpha = \bar{F}(0)$ , and  $\gamma \geq 1$ . The obtained probability model will be called the (mean scaled) *Gaussian exponential stop-loss model*. Making use of the standard normal distribution  $\Phi(x)$  with density  $\phi(x) = \Phi'(x)$ , the Gaussian component of (3.4) can be rewritten after the usual completion of squares as

$$(3.5) \quad \pi_G(x) = \mu \cdot \frac{\phi \left( \gamma + \frac{\alpha}{\gamma} \cdot \frac{x}{\mu} \right)}{\phi(\gamma)}, \quad x \geq 0.$$

It is worthwhile to mention that the limiting Gaussian component  $\gamma \rightarrow \infty$  of (3.4) is a mean scaled compound geometric exponential model (e.g. Panjer and Willmot(1992), Example 2.9.1 and Kaas et al.(1994), p. 39). Alternatively, the last model is known to be compound Poisson (e.g. Hürlimann(1990), Section 4, and Panjer and Willmot(1992), Example 1.3.4), and has been used in actuarial

risk theory by Willmot(1989), Boogaert and De Waegenere(1990), and Hürli-  
mann(1994).

#### 4. CONDITIONS FOR A FEASIBLE COMPOUND POISSON APPROXIMATION

To relate the parameter  $\gamma$  with the *volatility*  $1 + k^2$ , where  $k = \frac{\sigma}{\mu}$  is the coefficient of variation, one needs the *second order stop-loss transform*  $\pi_2(x) = E[(X - x)_+^2] = 2 \cdot \int_x^\infty \pi(t)dt$ . The required formulas for the Gaussian and exponential models are given by

$$(4.1) \quad \begin{aligned} \pi_{2,G}(x) &= 2\mu^2 \frac{\gamma}{\alpha} \cdot \frac{\overline{\Phi}\left(\gamma + \frac{\alpha}{\gamma} \cdot \frac{x}{\mu}\right)}{\varphi(\gamma)}, \quad x \geq 0, \\ \pi_{2,E}(x) &= 2\pi(x_0)m(x_0) \cdot \exp\left\{-\left(\frac{x-x_0}{m(x_0)}\right)\right\}, \quad x \geq x_0, \end{aligned}$$

where  $\overline{\Phi}(x) = 1 - \Phi(x)$ . Since the second order moment about the origin satisfies the relationships  $E[X^2] = (1 + k^2)\mu^2 = \pi^2(0)$ , one obtains through calculation of the decomposition  $\pi_2(0) = \{\pi_{2,G}(0) - \pi_{2,G}(x_0)\} + \pi_{2,E}(x_0)$  the equation

$$(4.2) \quad 2 \frac{\gamma}{\alpha\varphi(\gamma)} \cdot \left\{ \Phi\left(\gamma + \frac{\alpha}{\gamma} z\right) - \Phi(\gamma) + \frac{\varphi\left(\gamma + \frac{\alpha}{\gamma} z\right)}{\gamma + \frac{\alpha}{\gamma} z} \right\} = 1 + k^2,$$

where one sets  $x_0 = z \cdot \mu$ . For fixed  $\alpha = \overline{F}(0)$ ,  $z \geq 0$ , what is the possible range of values for the coefficient of variation when  $\gamma \geq 1$ ? It suffices to determine the values taken by the bivariate function

$$(4.3) \quad a(x, y) := x \cdot \exp\left\{\frac{1}{2}x^2\right\} \cdot \left\{ \Phi\left(x + \frac{y}{x}\right) - \Phi(x) + \frac{\varphi\left(x + \frac{y}{x}\right)}{x + \frac{y}{x}} \right\}, \quad x \geq 1, \quad y \geq 0.$$

For fixed  $x \geq 1$ , the curly bracket is monotone decreasing in  $y$ , hence its minimum is attained as  $y \rightarrow \infty$ . Therefore, to find the minimum of  $a(x, y)$ , it suffices to determine the minimum of

$$(4.4) \quad b(x) := \lim_{y \rightarrow \infty} a(x, y) = xe^{\frac{1}{2}x^2} \overline{\Phi}(x), \quad x \geq 1.$$

Since  $b'(x) = \exp\left\{\frac{1}{2}x^2\right\} \cdot \{(1 + x^2)\overline{\Phi}(x) - x\varphi(x)\}$  is non-negative, the function  $b(x)$  is increasing. For  $x \geq 1$  the function  $b(x)$ , hence also  $a(x, y)$ , takes values in the interval  $[\sqrt{e} \cdot \overline{\Phi}(1), \infty)$ . From (4.2) it follows that the volatility of the Gaussian exponential stop-loss model must satisfy the inequality

$$(4.5) \quad 1 + k^2 \geq \frac{2\sqrt{2\pi e} \cdot \bar{\Phi}(1)}{\alpha} = \frac{1.3117}{\alpha}.$$

For the typical case of a compound Poisson model, the volatility is given by

$$(4.6) \quad 1 + k^2 = 1 + \frac{1}{\lambda}(1 + v),$$

where  $\lambda$  is the Poisson parameter, and  $1 + v = \frac{m_2}{m_1^2}$  is the volatility of the severity distribution, with  $m_1, m_2$  the first two moments about the origin. Inserting (4.6) into (4.5) and noting that  $\alpha = \bar{F}(0) = 1 - e^{-\lambda}$ , the Gaussian exponential approximation is feasible provided

$$(4.7) \quad (1.3117) \left( \frac{\lambda}{1 - e^{-\lambda}} \right) - \lambda \leq 1 + v.$$

The function on the left hand side first decreases and then increases in  $\lambda$ . For fixed  $v$  there is a limited range of values  $\lambda \in [\lambda_{\min}(v), \lambda_{\max}(v)]$  for which (4.7) is feasible. If  $v \geq 0.75$  one has approximately  $[\lambda_{\min}(v), \lambda_{\max}(v)] = \left[0, \frac{1+v}{0.3117}\right]$ . For fixed severity volatility only “small” portfolios with  $\lambda < \lambda_{\max}(v)$  are feasible. Similarly, for fixed  $\lambda$ , only “highly volatile” portfolios for which  $v$  satisfies (4.7) are feasible.

## 5. SEMIPARAMETRIC DETERMINATION OF THE THRESHOLD

Since the Gaussian mean residual life function  $m_G(x) \rightarrow 0$  as  $x \rightarrow \infty$ , but  $m(x) \rightarrow$  positive constant for compound Poisson distributions under a regularity condition, the Gaussian stop-loss transform will underestimate the true value for large deductibles. To be effectively useful, a precise value must be assigned to the threshold  $x_0 = z \cdot \mu$  in (3.3), above which a simple exponential model is chosen as an approximation. Consider the case of given moments  $m_1, m_2$  and maximal amount  $b$  of the severity, which should be the most often encountered situation in practice. Then we choose the threshold in such a way that the unexplained Gaussian mean excess value in the infinite upper layer above the threshold takes the value of some “highly dangerous” compound Poisson distribution, where dangerousness is understood in terms of the stop-loss order or increasing convex order. The described semiparametric approach, for which only three (two when  $b \rightarrow \infty$ ) parameters need to be determined, allows for a computation of the threshold as follows.

Let  $X_G$  be a random variable with survival function  $\bar{F}_G(x)$  equal to the Gaussian component of (3.3). From the compound Poisson distribution of the random variable  $X$ , we suppose that in addition to the Poisson parameter  $\lambda$ , one knows the mean  $m_1$ , the relative variance  $v = \frac{m_2}{m_1^2} - 1$  of  $Y$ , and the relative maximal severity  $v_0 = \frac{b - m_1}{m_1}$ , where  $b$  is the maximal amount of the severity. Note that  $v_0$  can be interpreted as the maximal relative variance for given mean

$m_1$  and range  $[0, b]$  of the severity. Furthermore let  $v_r = \frac{v}{v_0}$  be the ratio of these relative variances.

In the following, the notation  $X \leq_{sl} Y$  means  $X$  precedes  $Y$  in stop-loss order by equal means, that is  $E[X] = E[Y]$  and  $\pi_X(x) \leq \pi_Y(x)$  for all  $x$ . Based on Hürlimann(1996/98), we construct a stop-loss upper bound

$$(5.1) \quad X = Y_1 + \dots + Y_N \leq_{sl} X^* = Y_1^* + \dots + Y_N^*,$$

where  $N$  is  $\text{Poisson}(\lambda)$  distributed, the  $Y_k^*$ 's are independent and identically distributed as  $Y^*$  and the  $Y_k^*$ 's are independent from  $N$ . According to the above approach, the threshold is now chosen such that the unexplained Gaussian conditional mean excess value in the infinite upper layer equals the corresponding distribution-free quantity obtained from the stop-loss upper bound  $X^*$ , in formulas:

$$(5.2) \quad m_G(x_0) = E[X_G - x_0 | X_G > x_0] = E[X^* - x_0 | X^* > x_0] = m^*(x_0).$$

Replacing the unexplained approximate component  $m_G(x_0)$  by a distribution-free quantity  $m^*(x_0)$  of the true compound Poisson amount  $m_X(x_0)$  intuitively reduces the modelling specification error, that is the possible wrong choice of an approximate model.

By (3.3) and (3.4), the left hand side of (5.2) equals

$$(5.3) \quad m_G(x_0) = m_G(z \cdot \mu) = \frac{\mu}{\alpha + \left(\frac{\alpha}{\gamma}\right)^2 z}.$$

Let us show how to evaluate the right hand side of (5.2). From Hürlimann(1996), Table 4, one first notes that  $Y^*$  is the random variable with support  $\{x_0^*, x_1^*, x_2^*, x_3^*\}$  and probabilities  $\{p_0, p_1, p_2, p_3\}$  such that

$$(5.4) \quad \begin{aligned} x_0^* &= 0, \quad x_1^* := \frac{x_1}{\mu} = \frac{1}{2} k^2, \quad x_2^* := \frac{x_2}{\mu} = \frac{1 + \frac{1}{2}(v_0 - v_r)}{\lambda}, \quad x_3^* := \frac{x_3}{\mu} = \frac{v_r}{v_r + v_0}, \\ p_0 &= \frac{v}{1+v}, \quad p_1 = \frac{v_0 - v}{(1+v)(1+v_0)}, \quad p_2 = \frac{v_0 - v}{(1+v_0)(v_r + v_0)}, \quad p_3 = \frac{v_r}{v_r + v_0}. \end{aligned}$$

Then the distribution-free mean residual life  $m^*(x_0) = \frac{\pi^*(x_0)}{F(x_0)}$  is described by the following (finite) analytical formulas (use (3.17) in Hürlimann(1996)):

$$(5.5) \quad \begin{aligned} \pi^*(x_0) &= \pi^*(z \cdot \mu) = \mu \cdot \left\{ 1 - z + \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n_1, n_2, n_3=0}^{\infty} p(n_1, n_2, n_3) \cdot \left(z - \sum_{i=1}^3 n_i x_i\right)_+ \right\}, \\ \bar{F}^*(x_0) &= \bar{F}^*(z \cdot \mu) = -\frac{d}{dx} \pi^*(x) \Big|_{x=x_0} = 1 - \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n_1, n_2, n_3=0}^{\infty} p(n_1, n_2, n_3) \cdot 1_{\left\{z > \sum_{i=1}^3 n_i x_i\right\}}, \\ p(n_1, n_2, n_3) &= \frac{(\lambda p_1)^{n_1} (\lambda p_2)^{n_2} (\lambda p_3)^{n_3}}{n_1! n_2! n_3!}. \end{aligned}$$



We remark that in the limiting case as  $v_0 \rightarrow \infty$  (i.e. an unbounded severity), one observes that  $x_2, x_3 \rightarrow \infty$ ,  $p_2, p_3 \rightarrow 0$  and (5.5) simplifies to

$$(5.6) \quad \begin{aligned} \pi^*(z \cdot \mu) &= \mu \cdot \left\{ 1 - z + \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{k^2}\right)^n \cdot \left(z - \frac{1}{2}nk^2\right)_+ \right\}, \\ \bar{F}^*(z \cdot \mu) &= 1 - \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{k^2}\right)^n \cdot 1_{\{z > \frac{1}{2}nk^2\}}, \end{aligned}$$

which shows by passing that  $\pi^*(z \cdot \mu)$  is approximately  $\frac{1}{2}\mu$  for sufficiently large  $z$ . The simple structure of (5.5) allows one to write

$$(5.7) \quad m^*(z \cdot \mu) = \mu \cdot \frac{1 - B_z - (1 - A_z)z}{1 - A_z} = \mu \cdot \left\{ \left( \frac{1 - B_z}{1 - A_z} \right) - z \right\},$$

where one has

$$(5.8) \quad A_z = \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n_1, n_2, n_3=0}^{\infty} p(n_1, n_2, n_3) \cdot 1_{\left\{z > \sum_{i=1}^3 n_i x_i\right\}},$$

$$(5.9) \quad B_z = \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n_1, n_2, n_3=0}^{\infty} \frac{\sum_{i=1}^3 n_i x_i}{p(n_1, n_2, n_3)} \cdot 1_{\left\{z > \sum_{i=1}^3 n_i x_i\right\}}.$$

In the limiting case  $v_0 \rightarrow \infty$  these expressions are replaced by

$$(5.10) \quad A_z = \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{k^2}\right)^n \cdot 1_{\left\{z > \frac{1}{2}nk^2\right\}},$$

$$(5.11) \quad B_z = \frac{1}{2} \exp\left\{-\frac{1}{k^2}\right\} \cdot \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \left(\frac{1}{k^2}\right)^{n-1} \cdot 1_{\left\{z > \frac{1}{2}nk^2\right\}}.$$

The solution  $x_0 = z \cdot \mu$  of (5.2), which uses the forms (5.3) and (5.7), is obtained from

$$(5.12) \quad \frac{1}{\alpha + \left(\frac{\alpha}{\gamma}\right)^2 z} = C_z - z, \quad C_z = \frac{1 - B_z}{1 - A_z},$$

as the solution of an implicit quadratic equation, and satisfies the formula

$$(5.13) \quad z = \frac{1}{2} \left\{ 1 - \frac{\gamma^2}{\alpha C_z} + \sqrt{\left(1 - \frac{\gamma^2}{\alpha C_z}\right)^2 + 4 \frac{\gamma^2}{\alpha} \left(1 - \frac{1}{\alpha C_z}\right)} \right\}.$$

As  $z$  increases, the quantities  $A_z, B_z$  increase and tend to the values  $A_z \rightarrow 1$  and  $B_z \rightarrow \frac{1}{2}$  as  $z \rightarrow \infty$  and  $v_0 \rightarrow \infty$ . In this situation, one has  $C_z \rightarrow \infty$  and there is a limiting maximal threshold, which is given by

$$(5.14) \quad z_{\max} = \frac{1}{2} \left\{ 1 + \sqrt{1 + 4 \frac{\gamma^2}{\alpha}} \right\}.$$

The values of  $z, \gamma$  are obtained iteratively by inserting (5.13) into the volatility equation (4.2). In our later examples the relations (5.13) and (5.14) often yield similar values of  $z, \gamma$ .

## 6. NUMERICAL IMPLEMENTATION AND ILLUSTRATIONS

The numerical evaluation of the unknown parameters  $z, \gamma$  simplifies if one takes into account the following property of the volatility equation (4.2). The well-known inequality  $x\bar{\Phi}(x) \leq \varphi(x)$  for  $x \geq 0$  shows that the true value of  $\gamma$  is bounded above by the solution  $\gamma_0$  of the pure Gaussian volatility equation

$$(6.1) \quad 2 \frac{\gamma_0}{\alpha} \frac{\bar{\Phi}(\gamma_0)}{\varphi(\gamma_0)} = 1 + k^2.$$

With this maximal value of  $\gamma$ , the corresponding threshold  $z$  is an implicit solution of (5.13). With the obtained value of  $z$ , it is possible to adjust for  $\gamma$  by using (4.2). The described iterative algorithm used to compute  $z, \gamma$  converges in our experience quickly to the accurate values.

Our first numerical illustration is based on the pension fund PK-230 of Held(1982), also discussed in Dufresne and Niederhauser(1997). A comparative study varying the relative maximal severity  $v_0$  is summarized in Tables 6.1 and 6.2. As a remarkable feature, the Gaussian exponential stop-loss model for  $v_0 = 8$ , which corresponds to the effective maximal severity of amount  $b = 487'000$ , is on the safe side from a reinsurance viewpoint. In general, increasing  $v_0$  up to  $v_0 = \infty$ , the stop-loss values do not always remain on the safe side. A very good fit in the right tail of the distribution is obtained for the limiting value  $v_0 = \infty$ .

Let us justify more formally the use of the simplest limiting model  $v_0 = \infty$ . This is done by comparing its quality of fit with that of the zero gamma, translated gamma and zero translated gamma approximations considered in Dufresne and Niederhauser(1997). The quality of fit is measured using 5 goodness-of-fit measures, which are “probability” analogues to those used in statistical estimation. Assume the data is grouped into  $m = 15$  classes with boundaries  $\xi_j = j \cdot \mu$ ,  $j = 0, \dots, m$ , and that the distribution function and stop-loss transform of the true compound Poisson distribution are known such that  $\hat{F}_j = \Pr(X \leq \xi_j)$ ,  $\hat{\pi}_j = E[(X - \xi_j)_+]$ ,  $j = 0, \dots, m$  are the observed data set. The corresponding approximated values are denoted by  $F_j, \pi_j, j = 0, \dots, m$ . We use the following goodness-of-fit measures:

### The Pearson's chi-square measure

$$\chi^2 = \sum_{j=1}^m \frac{(\hat{F}_j - \hat{F}_{j-1} - (F_j - F_{j-1}))^2}{F_j - F_{j-1}}$$

TABLE 6.1  
PARAMETERS OF THE GAUSSIAN EXPONENTIAL STOP-LOSS MODEL

Basic portfolio parameters					
expected number of claims	: $\lambda = 1.23$				
probability at least one claim occurs	: $\alpha = 0.70771$				
expected aggregate claims	: $\mu = 67'000$				
coefficient of variation of aggregate claims	: $k = 1.2737$				
relative variance of the claim size	: $v = \lambda k^2 - 1 = 0.99544$				
Calculation of the threshold: parameters and values					
$v_0$	$x_1$ $p_1$	$x_2$ $p_2$	$x_3$ $p_3$	$\gamma$	$z$
8	0.81116 0.39003	4.01446 0.0958	7.31707 0.01532	3.12823	2.37462
12	0.81116 0.42422	5.65734 0.07006	10.56911 0.0068653	3.21131	2.86238
16	0.81116 0.44232	7.29178 0.05495	13.82114 0.0038734	3.24389	3.1375
32	0.81116 0.47084	13.80849 0.02933	26.82927 0.00097117	3.28944	3.70555
$\infty$	0.81116 0.50114	$\infty$ 0	$\infty$ 0	3.32231	4.47972

TABLE 6.2  
STOP-LOSS TRANSFORM APPROXIMATIONS FOR THE PENSION FUND PK-230

Deductible	Compound	Gaussian exponential model				
$x \cdot \mu$	Poisson model	$v_0 = 8$	12	16	32	$\infty$
$1 \cdot \mu$	32000	32182	32224	32239	32260	32275
$2 \cdot \mu$	14308	14686	14763	14792	14831	14858
$3 \cdot \mu$	6368	6432	6446	6471	6510	6536
$4 \cdot \mu$	2762	2807	2764	2748	2734	2748
$5 \cdot \mu$	1186	1225	1185	1166	1135	1111
$6 \cdot \mu$	500	535	508	495	471	447
$7 \cdot \mu$	198	233	218	210	196	180
$8 \cdot \mu$	78	102	94	89	81	72
$9 \cdot \mu$	30	44	40	38	34	29
$10 \cdot \mu$	12	19	17	16	14	12
$11 \cdot \mu$	4	8	7	7	6	5
$12 \cdot \mu$	2	4	3	3	2	2
$13 \cdot \mu$	1	2	1	1	1	1
$14 \cdot \mu$	0	1	1	1	0	0
$15 \cdot \mu$	0	1	0	0	0	0

In statistics, this is the standard goodness-of-fit measure with grouped data, for which an elegant statistical theory exists (e.g. Hogg and Klugman(1984), p. 107).

**The weighted Cramér-von Mises minimum measure**

$$K = \sum_{j=0}^m \frac{m}{F_j \cdot (1 - F_j)} \cdot (F_j - \hat{F}_j)^2$$

The chosen weights are largest at the ends of the distribution, which render the K-statistic especially useful for the empirical testing of the tails of the distribution (e.g. Hogg and Klugman(1984), p. 135).

**The stop-loss distance measure**

$$SLD = \sum_{j=1}^m \left( \frac{\pi_j - \hat{\pi}_j}{\pi_j} \right)^2$$

**The limited expected value distance measure**

$$LED = \sum_{j=1}^m \left( \frac{LE_j - L\hat{E}_j}{LE_j} \right)^2, \text{ with } LE_j = \pi_0 - \pi_j, \quad L\hat{E}_j = \hat{\pi}_0 - \hat{\pi}_j$$

**The mean excess distance measure**

$$MED = \sum_{j=1}^m \left( \frac{ME_j - M\hat{E}_j}{ME_j} \right)^2, \text{ with } ME_j = \frac{\pi_j}{1 - F_j}, \quad M\hat{E}_j = \frac{\hat{\pi}_j}{1 - \hat{F}_j}$$

The last two “ad-hoc” measures have been used and motivated by some actuaries (Hogg and Klugman(1984), Klugman et al.(1998)). Apart from Carriere (1992) who studies the LED, it seems that these measures have not received too much attention. The goodness-of-fit analysis is summarized in Table 6.3. The analytical approximations are calculated using the parameter values  $\lambda = 1.23$

TABLE 6.3  
GOODNESS-OF-FIT COMPARISONS

Distribution	$\chi^2$	K	SLD	LED	MED
Gaussian exponential	0.00535	0.0234	0.33631	0.00068	0.88545
Zero gamma	0.00713	0.03917	3.15515	0.00094	0.27553
Translated gamma	0.02195	0.7894	2.71291	0.03076	0.25867
Zero translated gamma	0.01147	0.08547	1.28125	0.0011	0.10766

(Poisson parameter),  $\mu = 67000$  (mean),  $k = 1.2737$  (coefficient of variation),  $\gamma_1 = 1.895$  (skewness).

With respect to the first 4 criteria, the Gaussian exponential approximation displays the best fit. Note that a high MED-value does not necessarily mean a lack of fit. As observed in another recent study, the best overall fits in a set of alternatives show often higher MED-values (e.g. Hürlimann(2001)). This observation seems to be related with the longer tails of the distributions, but no precise explanation has been found so far for this phenomenon. Furthermore, a final comment on statistical estimation is in order. Based on the parameter vector  $(\lambda, \mu, k)$ , which is a reasonable choice in statistical studies, only the three parameter Gaussian exponential and zero gamma could actually be compared. The other distributions require knowledge of the skewness parameter, which is more difficult to estimate.

As a second case study, it is instructive to analyse the performance of the Gaussian exponential approximation by varying the individual claim size distribution. Tables 6.4 and 6.5 illustrate this issue using gamma and Pareto claim size distributions. These are prototypes for light tailed and heavy tailed distributions used in life insurance and reinsurance. The compound Poisson distributions are evaluated using a discretization of the continuous claim sizes and then applying Panjer's(1981) recursion.

TABLE 6.4  
STOP-LOSS TRANSFORMS OF COMPOUND POISSON GAMMA AND GAUSSIAN EXPONENTIAL

Deductible	Parameters			
	$\lambda = 5$	$\gamma = 5.39453$	$\lambda = 10$	$\gamma = 1.30378$
	$\mu = 103.09079$	$z = 5.93584$	$\mu = 206.18158$	$z = 1.89113$
	$k = 0.97497$		$k = 0.68941$	
	comp. Poisson	Gaussian exp.	comp. Poisson	Gaussian exp.
0	103.091	103.091	206.182	206.182
60	57.848	57.500	148.966	150.334
120	31.798	31.705	101.881	104.287
180	17.185	17.282	66.835	68.828
240	9.167	9.313	42.435	43.219
300	4.841	4.961	26.246	25.819
360	2.535	2.613	15.886	14.675
420	1.318	1.360	9.444	7.985
480	0.682	0.700	5.529	4.318
540	0.351	0.356	3.195	2.335
600	0.180	0.179	1.825	1.263
660	0.092	0.089	1.032	0.683
720	0.047	0.045	0.578	0.369
780	0.024	0.022	0.321	0.200
840	0.012	0.011	0.177	0.108
900	0.006	0.005	0.097	0.058

TABLE 6.5  
STOP-LOSS TRANSFORMS OF COMPOUND POISSON PARETO AND GAUSSIAN EXPONENTIAL

	Parameters			
Deductible	$\lambda = 5$ $\mu = 102.20617$ $k = 0.89853$ comp. Poisson	$\gamma = 2.67528$ $z = 3.23044$ Gaussian exp.	$\lambda = 10$ $\mu = 204.41235$ $k = 0.63356$ comp. Poisson	$\gamma = 1.15366$ $z = 1.75426$ Gaussian exp.
0	102.206	102.206	204.412	204.412
60	52.149	55.709	145.338	147.564
120	25.239	28.957	94.104	99.849
180	12.853	14.353	57.274	63.328
240	7.119	6.784	34.115	37.647
300	4.290	3.058	20.521	20.978
360	2.775	1.322	12.712	10.957
420	1.897	0.568	8.183	5.549
480	1.349	0.244	5.481	2.810
540	0.987	0.105	3.806	1.423
600	0.734	0.045	2.722	0.721
660	0.552	0.019	1.992	0.365
720	0.414	0.008	1.479	0.185
780	0.309	0.004	1.107	0.094
840	0.225	0.002	0.829	0.047
900	0.158	0.001	0.615	0.024

The Gaussian exponential approximation is excellent and usually yields “safe” net stop-loss premiums for deductibles below the threshold  $x_0 = z \cdot \mu$ . For values in the right tail beyond the threshold, the approximation seems only good for a light tailed claim size distribution. To improve the goodness-of-fit for heavy tailed claim sizes, we suggest replacing the exponential right tail by a Pareto right tail. A possible similar construction of the Gaussian Pareto model is left as a challenge to the interested practitioner.

ACKNOWLEDGEMENTS

The author is grateful to the referees for corrections and insightful comments.

REFERENCES

BARLOW, R.E. and PROSCHAN, F. (1975) *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, New York.  
BEARD, R.E., PENTIKÄINEN, T. and PESONEN, E. (1969/77/84) *Risk Theory. The Stochastic Basis of Insurance*. Methuen (1969). Chapman and Hall (1977/84).  
BENKTANDER, G. (1970) Schadenverteilung nach Grösse in der Nicht-Leben-Versicherung. *Bulletin of the Swiss Association of Actuaries*, 263-283.

- BHATTACHARJEE, M.C. (1982) The class of mean residual lives and some consequences. *SIAM Journal of Algebraic and Discrete Methods* **3**(1), 56-65.
- BOOGAERT, P. and DE WAEGENERE, A. (1990) Macro-economic version of a classical formula in risk theory. *Insurance: Mathematics and Economics* **9**, 155-162.
- BRYSON, M.C. and SIDDIQUI, M.M. (1969) Some criteria for aging. *Journal of the American Statistical Association* **64**, 1472-1483.
- CARRIERE, J. (1992) Limited expected value comparison tests. *Statistics and Probability Letters* **15**, 321-27.
- DAVIS, H.T. and FELDSTEIN, M.L. (1979) The generalized Pareto law as a model for progressively censored survival data. *Biometrika* **66**(2), 299-306.
- DUFRESNE, F. and NIEDERHAUSER, E. (1997) Some analytical approximations of stop-loss premiums. *Bulletin of the Swiss Association of Actuaries*, 25-47.
- EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, TH. (1997) Modelling Extremal Events for Insurance and Finance. *Applications of Mathematics – Stochastic Modelling and Applied Probability*, vol. 33. Springer.
- HELD, R. (1982) Zur rekursiven Berechnung von Stop-Loss Prämien für Pensionskassen. *Bulletin of the Swiss Association of Actuaries*, 67-89.
- HOGG, R. and KLUGMAN, S. (1984) *Loss Distributions*. John Wiley, New York.
- HÜRLIMANN, W. (1990) Pseudo compound Poisson distributions in risk theory. *ASTIN Bulletin* **20**, 57-79.
- HÜRLIMANN, W. (1994) On stable insurance business models. *25th ASTIN Colloquium*, Cannes.
- HÜRLIMANN, W. (1996) Improved analytical bounds for some risk quantities. *ASTIN Bulletin* **26**(2), 185-99.
- HÜRLIMANN, W. (1998) Extremal Moment Methods and Stochastic Orders. Application in Actuarial Science. Manuscript (available from the author).
- HÜRLIMANN, W. (2000a) Higher-degree stop-loss transforms and stochastic orders (I) theory. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* **XXIV**(3), 449-463.
- HÜRLIMANN, W. (2000b) Higher-degree stop-loss transforms and stochastic orders (II) applications. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* **XXIV**(3), 465-476.
- HÜRLIMANN, W. (2001) Financial data analysis with two symmetric distributions. First Prize in the Gunnar Benktander ASTIN Award Competition. *ASTIN Bulletin* **31**, 187-211.
- KAAS, R., VAN HEERWAARDEN, A.E. and GOOVAERTS, M.J. (1994) Ordering of Actuarial Risks. *CAIRE Education Series* **1**, Brussels.
- KODLIN, D. (1967) A new response time distribution. *Biometrics* **23**, 227-39.
- KLUGMAN, S., PANJER, H.H. and WILLMOT, G.E. (1998) *Loss Models. From Data to Decisions*. John Wiley, New York.
- MUTH, E.J. (1977) Reliability models with positive memory derived from the mean residual life function. In Tsokas, C.P. and Shimi, I.N. (Editors). *The Theory and Applications of Reliability*, vol. II, 401-434. Academic Press, New York.
- PANJER, H.H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bulletin* **12**, 22-26.
- PANJER, H.H. and WILLMOT, G.E. (1992) *Insurance Risk Models*. Society of Actuaries, Schaumburg, Illinois.
- WILLMOT, G.E. (1989) The total claims distribution under inflationary conditions. *Scandinavian Actuarial Journal*, 1-12.

WERNER HÜRLIMANN  
*Winterthur Life and Pensions  
 Value and Risk Management  
 Postfach 300  
 CH-8401 Winterthur  
 Switzerland*